An example of computation of the density of ones in probabilistic cellular automata by direct recursion

Henryk Fukś

Abstract We present a method for computing probability of occurence of 1s in a configuration obtained by iteration of a probabilistic cellular automata (PCA), starting from a random initial configuration. If the PCA is sufficiently simple, one can construct a set of words (or blocks of symbols) which is complete, meaning that probabilities of occurence of words from this set can be expressed as linear combinations of probabilities of occurence of these words at the previous time step. One can then setup and solve a recursion for block probabilities. We demonstrate an example of such PCA, which can be viewed as a simple model of diffusion of information or spread of rumors. Expressions for the density of ones are obtained for this rule using the proposed method.

1 Introduction

Binary probabilistic cellular automata (PCA) in one dimension are one of the most frequently studied types of cellular automata, and one of the most natural and most frequently encountered problems in PCA is what the author proposes to call *the density response problem*: If the proportion of ones in the initial configuration drawn from a Bernoulli distribution is ρ_0, what is the expected proportion of ones after t iterations of the PCA rule?

Of course, one could ask a similar question about the probability of occurence of longer blocks of symbols after t iterations of the PCA rule. Due to the complexity of PCA dynamics, it is clear that questions of this type are rather hopeless if one wants to know the answer for an arbitrary rule. In spite of this, it may still be possible to provide the answer if the rule is sufficiently simple.

Henryk Fukś

Department of Mathematics and Statistics, Brock University, St. Catharines, Canada. e-mail: hfuks@brocku.ca

One of the methods which can be used to do this is studying the structure of preimages of short blocks and detecting patterns present in them. This approach has been successfully used by the author for a number of deterministic CA rules, such as elementary rules 172, 142, 130 (references [4], [3], and [6] respectively), and several others. It has also been used for a special class of PCA known as single-transition α-asynchronous rules [7].

In this chapter, however, we would like to describe yet another method of computing probabilities of blocks of symbols, by setting up a system of recursive equations which can then be explicitly solved. Such a recursive system can be easily constructed for any rule for probabilities of *all* blocks, but it is normally too big and to complex to be solved. In certain cases, however, one can find a smaller set of blocks for which the recursion is solvable. We will present one such example, using a PCA which can be viewed as a simple model of diffusion of innovations or spread of rumors.

To give the reader a flavour of what to expect, let us informally define the aforementioned PCA rule. Suppose we have an infinite one-dimensional lattice where each site is occupied by an individual who has already adopted the innovation (1) or who has not adopted it yet (0). Initially the proportion of adopters is ρ_0. Once the individual adopts the innovation, he remains in state 1 forever. Individuals in state 0 change can their states to 1 (adopt the innovation) with probabilities depending on the state of nearest neighbours: if only the right (resp., left) neighbour already adopted, the probability is p (resp., q), and if both of them already adopted, the probability is r. What is the proportion of adopters ρ_t of after t iterations of the rule, assuming that the initial configuration is drawn from a Bernoulli distribution? We will show that the explicit formula for ρ_t can derived,

$$\rho_t = \begin{cases} 1 - E\left((\rho_0 q - 1)(\rho_0 p - 1)\right)^t - F(1-r)^t & \text{if } pq\rho_0^2 - (p+q)\rho_0 + r \neq 0, \\ 1 - (G + Ht)(1-r)^{t-1} & \text{if } pq\rho_0^2 - (p+q)\rho_0 + r = 0, \end{cases}$$

where E, F, G, H are constants depending on parameters p, q, r and ρ_0.

In order to accomplish this, we will start from some general theoretical remarks, considering PCA as maps in the space of shift-invariant probability measures, similarly as done in [10], [11], [9], and other works. More precisely, we will look at orbits of uniform Bernoulli measures under the action of PCA.

2 Probabilistic cellular automata

Probabilistic CA are often defined as stochastic dynamical systems. In this article, we will concentrate on Boolean CA in one dimension. Let $s_i(t)$ denote the state of the lattice site i at time t, where $i \in \mathbb{Z}$, $t \in \mathbb{N}$. We will further assume that $s_i(t) \in \{0,1\}$ and we will say that the site i is occupied (empty) at time t if $s_i(t) = 1$ (resp., $s_i(t) = 0$).

In a probabilistic cellular automaton, lattice sites simultaneously change states from 0 to 1 or from 1 to 0 with probabilities depending on states of local neighbours. A common method for defining PCA is to specify a set of local transition probabilities. For example, in order to define a nearest-neighbour PCA one has to specify the probability $w(s_i(t+1))|s_{i-1}(t), s_i(t), s_{i+1}(t))$ that the site $s_i(t)$ with nearest neighbors $s_{i-1}(t), s_{i+1}(t)$ changes its state to $s_i(t+1)$ in a single time step.

A more formal definition of nearest-neighbour PCA can be constructed as follows. Let r be a positive integer, called *radius of PCA*, and let us consider a set of independent Boolean random variables $X_{i,\mathbf{b}}$, where $i \in \mathbb{Z}$ and $\mathbf{b} \in \{0,1\}^{2r+1}$. Probability that the random variable $X_{i,\mathbf{b}}$ takes the value $a \in \{0,1\}$ will be assumed to be independent of i and denoted by $w(a|\mathbf{b})$,

$$Pr(X_{i,\mathbf{b}} = a) = w(a|\mathbf{b}). \tag{1}$$

Obviously, $w(1|\mathbf{b}) + w(0|\mathbf{b}) = 1$ for all $\mathbf{b} \in \{0,1\}^{2r+1}$. The update rule for the PCA is then defined by

$$s_i(t+1) = X_{i,\{s_{i-r}(t),\ldots,s_i(t),\ldots,s_{i+r}(t)\}}. \tag{2}$$

Note that new random variables X are used at each time step t, that is, random variables X used at the time step t are independent of those used at previous time steps.

Having the above definition in mind, we note that in order to fully define a nearest-neighbour PCA rule (i.e., rule with $r = 1$), it is enough to specify eight transition probabilities $w(1|x_1 x_2 x_3)$ for all $x_1, x_2, x_3 \in \{0,1\}$. Remaining eight probabilities, $w(0|x_1 x_2 x_3)$, can be obtained by $w(0|x_1 x_2 x_3) = 1 - w(1|x_1 x_2 x_3)$.

In any dynamical system, the main object of interest is the orbit of the system starting from a given initial point, and properties of this orbit. In the case of PCA, we often assume that the initial condition is "random" or "disordered", typically meaning that each $s_i(0)$ is set to 1 with a given probability ρ_0 and to 0 with probability $1 - \rho_o$, independently of each other. We then want to answer question of the type "After t iterations, what is the proportion of sites in state 1?" or "After t iterations, what is the probability of finding a pair of adjacent zeros"? In order to pose and answer questions of this kind rigorously, we will present an alternative definition of PCA, as maps in the space of probability measures.

2.1 Orbits of probability measures

Let $\mathscr{A} = \{0,1\}$ and $X = \mathscr{A}^{\mathbb{Z}}$. A finite sequence of elements of $\mathscr{A}$, $\mathbf{b} = b_1 b_2 \ldots, b_n$ will be called a *block* (or *word*) of length n. Set of all blocks of elements of $\mathscr{A}$ of all possible lengths will be denoted by $\mathscr{A}^{\star}$.

A *cylinder set* generated by the block $\mathbf{b} = b_1 b_2 \ldots, b_n$ and anchored at i is defined as

$$[\mathbf{b}]_i = \{\mathbf{x} \in \mathscr{A}^{\mathbb{Z}} : \mathbf{x}_{[i,i+n)} = \mathbf{b}\}. \tag{3}$$

The set of probability measures on the σ-algebra generated by cylinder sets of X will be denoted by $\mathfrak{M}(X)$. Details of construction of such measures, using Hahn-Kolmogorov theorem, can be found in [5]. These details, however, are not essential for our subsequent considerations. Given a probability measure $\mu \in \mathfrak{M}(X)$, measure of a cylinder set $[\mathbf{b}]_i$, denoted by $\mu([\mathbf{b}]_i)$, is often informally called a "probability of occurence of block $\mathbf{b}$ at site i".

Let the function $w : \mathscr{A} \times \mathscr{A}^{2r+1} \to [0,1]$, whose values are denoted by $w(a|\mathbf{b})$ for $a \in \mathscr{A}$, $\mathbf{b} \in \mathscr{A}^{2r+1}$, satisfying $\sum_{a \in \mathscr{A}} w(a|\mathbf{b}) = 1$, be called *local transition function* of *radius r*, and let its values be called *local transition probabilities*. *A probabilistic cellular automaton* with local transition function w is a map $F : \mathfrak{M}(X) \to \mathfrak{M}(X)$ defined as

$$(F\mu)([\mathbf{a}]_i) = \sum_{\mathbf{b} \in \mathscr{A}^{|\mathbf{a}|+2r}} w(\mathbf{a}|\mathbf{b})\mu([\mathbf{b}]_{i-r}) \text{ for all } i \in \mathbb{Z}, \mathbf{a} \in \mathscr{A}^{\star}, \tag{4}$$

where we define

$$w(\mathbf{a}|\mathbf{b}) = \prod_{j=1}^{|\mathbf{a}|} w(a_j|b_j b_{j+1} \dots b_{j+2r}). \tag{5}$$

When the function w takes values in the set $\{0,1\}$, the corresponding cellular automaton is called a *deterministic CA*.

In this paper, we will exclusively deal with shift-invariant probability measures for which $\mu([\mathbf{b}]_i)$ is independent of i. We will, therefore, drop the index i and simply write $\mu([\mathbf{b}])$. Moreover, we will be interested in orbits of Bernoulli measures ν_λ defined for $\lambda \in [0,1]$ by

$$\nu_\lambda([\mathbf{b}]) = \lambda^{\#1(\mathbf{b})}(1-\lambda)^{\#0(\mathbf{b})} \text{ for any } \mathbf{b} \in \mathscr{A}^{\star}, \tag{6}$$

where $\#0(\mathbf{b})$ and $\#1(\mathbf{b})$ denote the number of zeros and ones in $\mathbf{b}$. In order to simplify the notation, we define

$$P_t(\mathbf{b}) = (F^t \nu_\lambda)([\mathbf{b}]), \tag{7}$$

which will be informally referred to as "probability of occurence of block $\mathbf{b}$ after t iterations of PCA rule F". With this notation, eq. (4) can be written as

$$P_{t+1}(\mathbf{a}) = \sum_{\mathbf{b} \in \mathscr{A}^{|\mathbf{a}|+2r}} w(\mathbf{a}|\mathbf{b})P_t(\mathbf{b}), \tag{8}$$

for any $\mathbf{a} \in \mathscr{A}^{\star}$ and $t \in \mathbb{N}$. We will furthermore define

$$P_0(\mathbf{a}) = \nu_\lambda([\mathbf{a}]) = \lambda^{\#1(\mathbf{a})}(1-\lambda)^{\#0(\mathbf{a})} \tag{9}$$

for any $\mathbf{a} \in \mathscr{A}^{\star}$.

Elements of $\mathscr{A}^{\star}$ can be enumerated in lexicographical order, and corresponding probabilities arranged in an infinite column vector

$$\mathbf{P}_t = (P_t(0), P_t(1), P_t(00), P_t(01), P_t(10), P_t(11), P_t(000)\ldots)^T. \quad (10)$$

Before we continue, note that not all these probabilities are independent. Due to additivity of measure, the following relationships, know as consistency conditions, are valid for any $\mathbf{a} \in \mathscr{A}^\star$,

$$P_t(\mathbf{a}) = P_t(\mathbf{a}0) + P_t(\mathbf{a}1) = P_t(0\mathbf{a}) + P_t(1\mathbf{a}). \quad (11)$$

These conditions will be frequently used in our subsequent considerations.

Since each $P_{t+1}(\mathbf{a})$, by the virtue of eq. (8), is a linear combination of a finite number of $P_t(\mathbf{b})$ values, we can write

$$\mathbf{P}_{t+1} = \mathbf{M}\mathbf{P}_t, \quad (12)$$

where the infinite matrix $\mathbf{M}$ is defined by eq. (8). This yields the following expression for probabilities of all finite words,

$$\mathbf{P}_t = \mathbf{M}^t \mathbf{P}_0, \quad (13)$$

where components of $\mathbf{P}_0$ are defined in eq. (9). In theory, the above equation gives us a complete solution of the problem of determining the orbit of Bernoulli measure under iterations of a PCA rule. In practice, however, computing powers of an infinite matrix is a daunting, if not impossible, task.

In practical applications, however, we rarely need *all* probabilities $P_t(\mathbf{a})$, that is, all components of the vector $\mathbf{P}_t$. Sometimes we are interested only in one specific probability, for example, $P_t(1)$. For a binary PCA, the expected value of a given lattice site after t iterations of the rule is equal to $1 \cdot P_t(1) + 0 \cdot P_t(0) = P_t(1)$, and for that reason, $P_t(1)$ is sometimes referred to as an expected *density of ones*, to be denoted by ρ_t,

$$\rho_t = P_t(1). \quad (14)$$

Note that for Bernoulli measure ν_λ we have $\rho_0 = \lambda$. Given ρ_0, could one find an explicit expression for ρ_t for a given PCA using eq. (12)? This problem will be called a *density response problem*. Although it cannot be solved in a general case, we will demonstrate that for a sufficiently simple PCA it is a doable task.

The idea is to setup a recursion similar to (12), but using a "smaller" set of block probabilities, for which the matrix $\mathbf{M}$ has somewhat simpler structure, lending itself to direct computation of $\mathbf{M}^t$. If we could then express ρ_t in terms of block probabilities from this "smaller" set, we would solve the density response problem.

Let us define the concept of the "smaller" set first. A set of words $\mathscr{A}^\star \supset C = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots\}$ will be called *complete* with respect to a PCA rule F if for every $\mathbf{a} \in C$ and $t \in \mathbb{N}$, $P_{t+1}(\mathbf{a})$ can be expressed as a linear combination of $P_t(\mathbf{a}_1), P_t(\mathbf{a}_2), P_t(\mathbf{a}_3), \ldots$. We will show a concrete example of a complete set in the next section.

3 Example PCA rule

As an example, we will consider a PCA rule which generalizes some of the CA rules investigated in [1]. This PCA can be viewed as a simple model for diffusion of innovations, spread of rumors, or a similar process involving transport of information between neighbours. We consider an infinite one-dimensional lattice where each site is occupied by an individual who has already adopted the innovation (1) or who has not adopted it yet (0). Once the individual adopts the innovation, he remains in state 1 forever. Individuals in state 0 change can their states to 1 (adopt the innovation) with probabilities depending on the state of nearest neighbours. All changes of states take place simultaneously. This process can be formally described as a radius 1 binary PCA with the following transition probabilities,

$$w(1|000) = 0, w(1|001) = p, w(1|010) = 1, w(1|011) = 1, \qquad (15)$$
$$w(1|100) = q, w(1|101) = r, w(1|110) = 1, w(1|111) = 1,$$

where p, q, r are fixed parameters of the model, $p, q, r \in [0, 1]$. In order to illustrate the difficulty of computing block probabilities for this rule, let us write eq. (8) for blocks **a** of length 1 and 2,

$$P_{t+1}(0) = P_t(000) + (1-p)P_t(001) + (1-q)P_t(100) + (1-r)P_t(101),$$
$$P_{t+1}(1) = pP_t(001) + P_t(010) + P_t(011) + qP_t(100) + rP_t(101) + P_t(110) + P_t(111),$$

$$P_{t+1}(00) = P_t(0000) + (1-p)P_t(0001) + (1-q)P_t(1000) + (1-p)(1-q)P_t(1001),$$
$$P_{t+1}(01) = pP_t(0001) + (1-p)P_t(0010) + (1-p)P_t(0011) + p(1-q)P_t(1001)$$
$$\qquad + (1-r)P_t(1010) + (1-r)P_t(1011),$$
$$P_{t+1}(10) = (1-q)P_t(0100) + (1-r)P_t(0101) + qP_t(1000) + (1-p)qP_t(1001)$$
$$\qquad + (1-q)P_t(1100) + (1-r)P_t(1101),$$
$$P_{t+1}(11) = pP_t(0010) + pP_t(0011) + qP_t(0100) + rP_t(0101) + P_t(0110) + P_t(0111)$$
$$\qquad + pqP_t(1001) + rP_t(1010) + rP_t(1011) + qP_t(1100)$$
$$\qquad + rP_t(1101) + P_t(1110) + P_t(1111).$$

As we can see, in order to know $P_{t+1}(1)$, we need to know probabilities of blocks of length 3 at time step t, and in order to compute these, we would need probabilities of blocks of length 5 at time step $t-2$, etc.

We will now show, however, that for the PCA rule defined in eq. (15) a complete subset of $\mathscr{A}^\star$ can be constructed. This subset consists of clusters of zeros bounded by 1 on each side, that is, of blocks of the type 10^n1, where $n \in \mathbb{N}$ and 0^n denotes n consecutive zeros.

Proposition 1. *The set $C = \{101, 1001, 100001, \ldots\}$ is complete with respect to the PCA rule defined in in eq. (15).*

In order to prove this, we need to show that every $P_{t+1}(10^n 1)$ can be expressed as a linear combination of probabilities of the type $P_t(10^k 1)$. Let us write eq. (8) for $\mathbf{a} = 10^n 1$. Two cases must be distinguished, $n = 1$ and $n > 1$. For $n = 1$, we have

$$
\begin{aligned}
P_{t+1}(101) = \; & p(1-q)P_t(01001) + (1-r)P_t(01010) + (1-r)P_t(01011) \\
& + pqP_t(10001) + (1-p)qP_t(10010) + (1-p)qP_t(10011) + p(1-q)P_t(11001) \\
& + (1-r)P_t(11010) + (1-r)P_t(11011). \quad (16)
\end{aligned}
$$

By consistency conditions, $P_t(10010) + P_t(10011) = P_t(1001)$ and $P_t(11010) + P_t(11011) = P_t(1101)$, as well as $P_t(01001) + P_t(11001) = P_t(1001)$. This yields

$$
\begin{aligned}
P_{t+1}(101) = \; & (1-r)P_t(01010) + (1-r)P_t(01011) + pqP_t(10001) \\
& + (1-p)qP_t(1001) + p(1-q)P_t(1001) + (1-r)P_t(1101), \quad (17)
\end{aligned}
$$

and further reduction is possible using $P_t(01010) + P_t(01011) = P_t(0101)$ and $P_t(0101) + P_t(1101) = P_t(101)$. The final result is

$$
P_{t+1}(101) = (1-r)P_t(101) + (p - 2pq + q)P_t(1001) + pqP_t(10001). \quad (18)
$$

For $n > 1$, using a similar procedure (omitted here), we obtain

$$
P_{t+1}(10^n 1) = (1-p)(1-q)P_t(10^n 1) + (p - 2pq + q)P_t(10^{n+1} 1) + pqP_t(10^{n+2} 1). \quad (19)
$$

Equations (18) and (19) clearly show that the set C is complete. $\square$

4 Calculations of $P_t(10^n 1)$

Having a complete set of block probabilities, we can now write eq. (18) and (19) in matrix form,

$$
\begin{bmatrix}
P_{t+1}(101) \\
P_{t+1}(1001) \\
\vdots \\
P_{t+1}(10^n 1) \\
\vdots
\end{bmatrix}
=
\begin{bmatrix}
\tilde{a} & b & c & 0 & 0 & 0 & \cdots \\
0 & a & b & c & 0 & 0 & \\
0 & 0 & a & b & c & 0 & \\
0 & 0 & 0 & a & b & c & \\
\vdots & & & & & & \ddots
\end{bmatrix}
\begin{bmatrix}
P_t(101) \\
P_t(1001) \\
\vdots \\
P_t(10^n 1) \\
\vdots
\end{bmatrix},
\quad (20)
$$

where $a = (1-p)(1-q)$, $\tilde{a} = 1 - r$, $b = p - 2pq + q$, and $c = pq$.

Let us define

$$\mathbf{M} = \begin{bmatrix} \tilde{a} & b & c & 0 & 0 & 0 & \cdots \\ 0 & a & b & c & 0 & 0 & \\ 0 & 0 & a & b & c & 0 & \\ 0 & 0 & 0 & a & b & c & \\ \vdots & & & & & \ddots & \end{bmatrix}, \quad \mathbf{P}_t = \begin{bmatrix} P_t(101) \\ P_t(1001) \\ \vdots \\ P_t(10^n 1) \\ \vdots \end{bmatrix}. \tag{21}$$

We will use $\mathrm{diag}(x_1, x_2, x_3, \ldots)$ to denote an infinite matrix with $x_1, x_2, x_3, \ldots$ on the diagonal and zeros elsewhere. Similarly, $\mathrm{sdiag}(x_1, x_2, x_3, \ldots)$ will denote shifted diagonal matrix having $x_1, x_2, x_3, \ldots$ on the line above the diagonal and zeros elsewhere, and $^2\mathrm{sdiag}(x_1, x_2, x_3, \ldots)$ will denote doubly-shifted diagonal matrix, with $x_1, x_2, x_3, \ldots$ on the second line above the diagonal and zeros elsewhere. With this notation, we have

$$\mathbf{M} = \mathbf{A} + \mathbf{B} + \mathbf{C}, \tag{22}$$

where

$$\mathbf{A} = \mathrm{diag}(\tilde{a}, a, a, \ldots), \tag{23}$$
$$\mathbf{B} = \mathrm{sdiag}(b, b, b, \ldots), \tag{24}$$
$$\mathbf{C} = {}^2\mathrm{sdiag}(c, c, c, \ldots). \tag{25}$$

Now,

$$\mathbf{P}_t = \mathbf{M}^t \mathbf{P}_0, \tag{26}$$

and we need to compute $\mathbf{M}^t$. We will do it by considering a special case first.

4.1 Special case: $\tilde{a} = a$

When $\tilde{a} = a$, matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ pairwise commute, thus we can use the trinomial expansion formula,

$$\mathbf{M}^t = (\mathbf{A} + \mathbf{B} + \mathbf{C})^t = \sum_{i+j+k=t} \binom{t}{i, j, k} \mathbf{A}^i \mathbf{B}^j \mathbf{C}^k, \tag{27}$$

where

$$\binom{t}{i, j, k} = \frac{t!}{i! j! k!}. \tag{28}$$

Generalizing the previously introduced notation, let

$${}^n\mathrm{sdiag}(x_1, x_2, x_3, \ldots) \tag{29}$$

denote n-times shifted diagonal matrix, which has $x_1, x_2, x_3, \ldots$ entries on the n-th line above the diagonal and zeros elsewhere. It is straightforward to prove that

$$\mathbf{A}^i = \mathrm{diag}(a^i, a^i, a^i, \ldots), \tag{30}$$

$$\mathbf{B}^j = {}^j\mathrm{sdiag}(b^j, b^j, b^j, \ldots), \tag{31}$$

$$\mathbf{C}^k = {}^{2k}\mathrm{sdiag}(c^k, c^k, c^k, \ldots), \tag{32}$$

and, consequently,

$$\mathbf{A}^i \mathbf{B}^j \mathbf{C}^k = {}^{j+2k}\,\mathrm{sdiag}(a^i b^j c^k, a^i b^j c^k, a^i b^j c^k, \ldots). \tag{33}$$

In the first row of the above matrix, the only non-zero element $(a^i b^j c^k)$ is in the column $1 + j + 2k$. In the second row, the only non-zero element $(a^i b^j c^k)$ is in the column $2 + j + 2k$, and so on. This means that

$$\mathbf{A}^i \mathbf{B}^j \mathbf{C}^k \mathbf{P}_0 = \begin{bmatrix} a^i b^j c^k P_0(10^{1+j+2k}1) \\ a^i b^j c^k P_0(10^{2+j+2k}1) \\ a^i b^j c^k P_0(10^{3+j+2k}1) \\ \vdots \end{bmatrix}. \tag{34}$$

Using the above and the fact that $P_0(10^n 1) = \rho_0^2(1-\rho_0)^n$, we can now write

$$\mathbf{P}_t = \mathbf{M}^t \mathbf{P}_0 = \sum_{i+j+k=t} \binom{t}{i,j,k} \begin{bmatrix} \tilde{a}^i b^j c^k \rho_0^2(1-\rho_0)^{1+j+2k} \\ a^i b^j c^k \rho_0^2(1-\rho_0)^{2+j+2k} \\ a^i b^j c^k \rho_0^2(1-\rho_0)^{3+j+2k} \\ \vdots \end{bmatrix}. \tag{35}$$

We finally obtain

$$P_t(10^l 1) = \sum_{i+j+k=t} \binom{t}{i,j,k} a^i b^j c^k \rho_0^2(1-\rho_0)^{l+j+2k}$$

$$= \rho_0^2(1-\rho_0)^l \sum_{i+j+k=t} \binom{t}{i,j,k} a^i [b(1-\rho_0)]^j [c(1-\rho_0)^2]^k$$

$$= \rho_0^2(1-\rho_0)^l \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t. \tag{36}$$

4.2 General case

We are now ready to handle the general case, without the $\tilde{a} = a$ assumption. Let us first note that t-th powers of matrices

$$\begin{bmatrix} \tilde{a}\ b\ c\ 0\ 0\ 0\ \cdots \\ 0\ a\ b\ c\ 0\ 0 \\ 0\ 0\ a\ b\ c\ 0 \\ 0\ 0\ 0\ a\ b\ c \\ \vdots \qquad\qquad \ddots \end{bmatrix}^t , \quad \begin{bmatrix} a\ b\ c\ 0\ 0\ 0\ \cdots \\ 0\ a\ b\ c\ 0\ 0 \\ 0\ 0\ a\ b\ c\ 0 \\ 0\ 0\ 0\ a\ b\ c \\ \vdots \qquad\qquad \ddots \end{bmatrix}^t \tag{37}$$

differ only in their first row. This implies that the expression for $P_t(10^l 1)$ given in eq. (36) remains valid for $l > 1$ even if $\tilde{a} \neq a$. We only need to consider $l = 1$ case, that is, to compute $P_t(101)$. This can be done by writing eq. (18) and replacing $P_t(1001)$ and $P_t(10001)$ by appropriate expressions obtained from eq. (36),

$$P_{t+1}(101) = \tilde{a}P_t(101) + b\rho_0^2(1-\rho_0)^2 \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t$$
$$+ c\rho_0^2(1-\rho_0)^3 \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t. \tag{38}$$

This can be written as

$$P_{t+1}(101) = \tilde{a}P_t(101) + K\theta^t, \tag{39}$$

where

$$K = b\rho_0^2(1-\rho_0)^2 + c\rho_0^2(1-\rho_0)^3, \tag{40}$$
$$\theta = a + b(1-\rho_0) + c(1-\rho_0)^2. \tag{41}$$

Eq. (39) is a first-order non-homogeneous difference equation for $P_t(101)$, and, as such, it can be easily solved by standard methods [2]. The solution is

$$P_t(101) = P_0(101)\tilde{a}^t + K \sum_{i=1}^{t} \tilde{a}^{t-i} \theta^{i-1}. \tag{42}$$

The sum on the right hand side is a partial sum of geometric series if $\tilde{a} \neq \theta$, or of an arithmetic series when $\tilde{a} = \theta$. Using appropriate formulae for partial sums of geometric and arithmetic series one obtains

$$P_t(101) = \begin{cases} P_0(101)\tilde{a}^t + K(\tilde{a}^t - \theta^t)/(\tilde{a} - \theta) & \text{if } \tilde{a} \neq \theta, \\ P_0(101)\tilde{a}^t + K\tilde{a}^{t-1}t & \text{if } \tilde{a} = \theta. \end{cases} \tag{43}$$

Taking $P_0(101) = \rho^2(1-\rho)$ and replacing K and θ by their definitions we obtain, for $a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2 \neq 0$ (which is equivalent to $\tilde{a} \neq \theta$),

$$P_t(101) = \frac{\rho_0^2(1-\rho_0)^2(b + c - c\rho_0)}{a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2} \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t$$
$$+ \frac{\rho_0^2(1-\rho_0)(a - \tilde{a})}{a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2} \tilde{a}^t. \tag{44}$$

For $a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2 = 0$ (equivalent to $\tilde{a} = \theta$), the solution is slightly simpler,

$$P_t(101) = \rho_0^2 (1-\rho_0) \left(\left(c\rho_0{}^2 - (b+2c)\rho_0 + b + c \right) t + \tilde{a} \right) \tilde{a}^{t-1}. \qquad (45)$$

We now have expressions for $P_t(10^l 1)$ for $l = 1$ (eq. 44 and 45) and for $l > 1$ (eq. 36).

5 Cluster expansion

We are finally ready to compute ρ_t. To do this, we will use the formula

$$P_t(0) = \sum_{k=1}^{\infty} k P_t(10^k 1), \qquad (46)$$

which we will refer to as "cluster expansion". Various proofs of this formula can be given (see, for example, [12]), but we will show here that it is a direct consequence of additivity of measure.

Consider a cylinder set of a single zero anchored at i, $[0]_i$. A single zero must belong to a cluster of zeros of size k with possible values of k varying from 1 to infinity. If it belongs to a cluster of k zeros, than it must be the j-th zero of the cluster, with possible values of j varying from 1 to k. Therefore,

$$[0]_i = \bigcup_{k=1}^{\infty} \bigcup_{j=1}^{k} [10^k 1]_{i-j}. \qquad (47)$$

Since all the cylinder sets on the right hand side are mutually disjoint, their measures add up, thus

$$(F^t \nu_\lambda)([0]_i) = \sum_{k=1}^{\infty} \sum_{j=1}^{k} (F^t \nu_\lambda)([10^k 1]_{i-j}). \qquad (48)$$

The measure is shift-invariant, thus $(F^t \nu_\lambda)([10^k 1]_{i-j}) = P_t(10^k 1)$, and we obtain

$$P_t(0) = \sum_{k=1}^{\infty} \sum_{j=1}^{k} P_t(10^k 1), \qquad (49)$$

which yields eq. (46), as desired.

We can now compute $P_t(0)$ using the cluster expansion formula and eq. (44), (45) and (36). We will first consider the case of $a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2 \neq 0$, that is, using eq. (44) for $P_t(101)$.

$$P_t(0) = \sum_{l=1}^{\infty} l P_t(10^l 1) = \frac{\rho_0^2 (1-\rho_0)^2 (b+c-c\rho_0)}{a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2} \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t$$

$$+ \frac{\rho_0^2 (1-\rho_0)(a-\tilde{a})}{a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2} \tilde{a}^t + \sum_{l=2}^{\infty} l \rho_0^2 (1-\rho_0)^l \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t.$$

$$(50)$$

Since

$$\sum_{l=2}^{\infty} l(1-\rho_0)^l = \frac{(1+\rho_0)(1-\rho_0)^2}{\rho_0^2}, \tag{51}$$

we obtain

$$P_t(0) = \frac{\rho_0^2 (1-\rho_0)^2 (b+c-c\rho_0)}{a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2} \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t$$

$$+ \frac{\rho_0^2 (1-\rho_0)(a-\tilde{a})}{a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2} \tilde{a}^t + (1+\rho_0)(1-\rho_0)^2 \left(a + b(1-\rho_0) + c(1-\rho_0)^2\right)^t.$$

$$(52)$$

After substitution of $\tilde{a}, a, b, c$ and simplification, as well as taking $\rho_t = 1 - P_t(0)$, the following expression for ρ_t is obtained,

$$\rho_t = 1 - \frac{(1-\rho_0)^2 (r - (p-r+q)\rho_0)}{pq\rho_0^2 - (p+q)\rho_0 + r} \left((\rho_0 q - 1)(\rho_0 p - 1)\right)^t$$

$$- \frac{\rho_0^2 (1-\rho_0)((q-1)p - q + r)}{pq\rho_0^2 - (p+q)\rho_0 + r} (1-r)^t. \quad (53)$$

When $a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2 = 0$, similar calculations can be performed, but this time using using eq. (45) for $P_t(101)$. After simplification, this yields

$$\rho_t = 1 - (1-\rho_0)\left(\rho_0^2 (\rho_0 - 1)(pq\rho_0 - p + pq - q)t + 1 - r\right)(1-r)^{t-1}. \quad (54)$$

Let us summarize this in a more readable form, noticing that after substitution of $\tilde{a}, a, c, b$ by their definitions the condition $a - \tilde{a} + b(1-\rho_0) + c(1-\rho_0)^2 = 0$ becomes $pq\rho_0^2 - (p+q)\rho_0 + r = 0$. Our final expression for the density of ones can be written as

$$\rho_t = \begin{cases} 1 - E\left((\rho_0 q - 1)(\rho_0 p - 1)\right)^t - F(1-r)^t & \text{if } pq\rho_0^2 - (p+q)\rho_0 + r \neq 0, \\ 1 - (G + Ht)(1-r)^{t-1} & \text{if } pq\rho_0^2 - (p+q)\rho_0 + r = 0, \end{cases}$$

$$(55)$$

where definitions of E, F, G, H can be figured out by comparing the above to eq. (52) and (53).

We can see that in the non-degenerate case (when $pq\rho_0^2 - (p+q)\rho_0 + r \neq 0$), the limit $\rho_\infty = \lim_{t \to \infty} \rho_t$ always exists, and that ρ_t approaches ρ_∞ exponentially fast, excluding special cases when $\rho_t = \text{const}$ (such as $\rho_0 = 0$ or $p = q = 0$, $r = 1$).

In the degenerate case, ρ_∞ always exists as well, but the approach of ρ_t to ρ_∞ is linearly-exponential.

It is worth noting that the existence of the degenerate case is a fairly subtle phenomenon, and that it would be very difficult to discover the linearly-exponential convergence by computer simulations alone. This illustrates the point that having a formula for ρ_t brings some advantages, and that the search for such formulae is worthwhile.

As a separate remark, let us note that deterministic CA are nothing else but special cases of PCA, thus we can choose integer values of p, q, r and obtain relevant expression for ρ_t for a number of elementary CA rules (ECA), as follows.
ECA rule 206 ($p = 1, q = 0, r = 0$) or rule 220 ($p = 0, q = 1, r = 0$)

$$\rho_t = 1 - \rho_0(1 - \rho_0) - (1 - \rho_0)^{t+2}, \tag{56}$$

ECA rule 222 ($p = q = 1, r = 0$)

$$\rho_t = 1 + \frac{2(1 - \rho_0)^{2t+1} + \rho_0(1 - \rho_0)}{\rho_0 - 2}, \tag{57}$$

ECA rule 236 ($p = 0, q = 0, r = 1$)

$$\rho_t = 1 - (\rho_0 + 1)(1 - \rho_0)^2, \tag{58}$$

ECA rule 238 ($p = 1, q = 0, r = 1$) or rule 252 ($p = 0, q = 1, r = 1$),

$$\rho_t = 1 - (1 - \rho_0)^{t+1}, \tag{59}$$

ECA rule 254 ($p = q = r = 1$)

$$\rho_t = 1 - (1 - \rho_0)^{2t+1}. \tag{60}$$

The above formulae agree with those derived informally in [1].

6 Conclusions

We presented a method for computing the density of ones in the orbit of the Bernoulli measure under the action of a probabilistic cellular automaton, using a simple PCA rule as an example. For this rule, we were able to construct a complete set of block probabilities, and then solve the resulting recurrence relations. By using the cluster expansion, we then obtained the required density of ones. Although this method is obviously applicable only to PCA rules with rather simple dynamics, it may be possible to find other PCA rules with complete sets, thus making the method useful for them. Generalization of the rule used in this paper to larger neighbourhood sizes comes to mind as a first possibility, and sufficiently simple deterministic

rules, such as asymptotic emulators of identity investigated in [8], are also possible candidates.

7 Acknowledgments

References

1. N. Boccara and H. Fukś. Modeling diffusion of innovations with probabilistic cellular automata. In M. Delorme and J. Mazoyer, editors, *Cellular Automata: A Parallel Model*, Dordrecht, 1998. Kluwer Academic Publishers.
2. P. Cull, M. Flahive, and R. Robson. *Difference equations*. Springer, 2004.
3. H. Fukś. Dynamics of the cellular automaton rule 142. *Complex Systems*, 16:123–138, 2006.
4. H. Fukś. Probabilistic initial value problem for cellular automaton rule 172. *DMTCS proc.*, AL:31–44, 2010.
5. H. Fukś. Construction of local structure maps for cellular automata. *J. of Cellular Automata*, 7:455–488, 2013.
6. H. Fukś and A. Skelton. Response curves for cellular automata in one and two dimensions – an example of rigorous calculations. *International Journal of Natural Computing Research*, 1:85–99, 2010.
7. H Fukś and A. Skelton. Orbits of Bernoulli measure in asynchronous cellular automata. *Dis. Math. Theor. Comp. Science*, AP:95–112, 2011.
8. H. Fukś and J-M. Gómez Soto. Exponential convergence to equilibrium in cellular automata asymptotically emulating identity. *Complex Systems*, 23:1–26, 2014.
9. P. Kůrka. On the measure attractor of a cellular automaton. *Discrete and Continuous Dynamical Systems*, pages 524 – 535, 2005.
10. Petr Kůrka and Alejandro Maass. Limit sets of cellular automata associated to probability measures. *Journal of Statistical Physics*, 100:1031–1047, 2000.
11. Marcus Pivato. Conservation laws in cellular automata. *Nonlinearity*, 15(6):1781, 2002.
12. D. Stauffer and A. Aharony. *Introduction to percolation theory*. Taylor and Francis, 1994.